



Sequence Listing

COPY OF PAPERS
ORIGINALLY FILED

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Gorman, Cornelia M.,
Groskreutz, Debyra J.

(ii) TITLE OF INVENTION: Prohormone Convertase Transformed Cells and
Polypeptide Synthesis

(iii) NUMBER OF SEQUENCES: 57

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 1 DNA Way
(C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WinPatin (Genentech)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/026143
(B) FILING DATE: 01-Mar-1993
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/887265
(B) FILING DATE: 22-MAY-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/803631
(B) FILING DATE: 06-DEC-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US92/10621
(B) FILING DATE: 04-DEC-1992

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Love, Richard B.
(B) REGISTRATION NUMBER: 34,659
(C) REFERENCE/DOCKET NUMBER: P0748P3

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650/225-5530
(B) TELEFAX: 650/952-9881

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2355 base pairs

(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCTAGATCTA GCTGGTGTGT CTCTGATCTT GCTTCTTTTC TCCCAGCCCT 50
TCCTACTTGT GTGAGAACAA GGTTTTGAGC CATGGAGCAA AGAGGTTGGA 100
CTCTGCAGTG TACTGCTTTC GCCTTCTTTT GCGTTTGGTG TGACTAAGC 150
AGTGTAAGAG CAAAGAGGCA GTTTGTAAAT GAATGGGCGG CGGAGATCCC 200
CGGAGGGCAA GAAGCTGCCT CTGCCATCGC CGAAGAACTG GGGTATGACC 250
TTTTGGGTCA GATTGGATCA CTTGAAAATC ACTATTTATT CAAACACAAA 300
AGCCATCCTC GGAGGTCCCG AAGAAGCGCT CTCATATCA CTAAGAGGTT 350
ATCTGATGAT GATCGTGTGA CGTGGGCTGA ACAACAGTAT GAAAAAGAGA 400
GAAGTAAACG TTCAGTTCAA AAAGACTCAG CATTGGATCT CTTCAATGAT 450
CCAATGTGGA ATCAGCAGTG GTACTTGCAA GATACCAGAA TGAATGCAGC 500
TCTGCCCAAG CTGGACCTTC ATGTAATACC TGTGTTGGGAA AAGGGTATTA 550
CTGGCAAAGG AGTTGTTATT ACTGTACTGG ATGATGGCTT GGAGTGAAT 600
CACACAGACA TTTATGCCAA TTATGATCCA GAGGCTAGCT ATGATTTTAA 650
CGATAATGAT CATGATCCAT TTCCCGATA TGATCTCACA AATGAAAACA 700
AACATGGAAC AAGATGTGCA GGTGAAATTG CCATGCAAGC AAATAATCAC 750
AAGTGTGGGG TTGGAGTTGC ATATAATTCC AAAGTTGGAG GCATAAGAAT 800
GCTGGATGGC ATTGTAAGT ATGCCATTGA GGCTAGTTCA ATTGGATTCA 850
ACCCTGGCCA TGTGGATATT TACAGTGCAA GCTGGGGCCC TAATGATGAT 900
GGAAAACTG TGGAGGGGCC TGGCAGACTA GCCCAGAAGG CATTGAATA 950
TGGTGTCAA CAGGGGAGAC AAGGGAAAGG CTCCATCTTT GTCTGGGCTT 1000
CAGGGAATGG GGGTCGTCAG GGAGATAACT GTGACTGTGA TGGCTACACA 1050
GACAGCATTT ACACCATCTC TATCAGCAGT GCCTCCCAGC AAGGCCTGTC 1100
ACCTTGGTAT GCAGAGAAGT GTTCTTCCAC ATTGGCTACC TCCTACAGCA 1150
GTGGTGATTA CACAGACCAG CGAATAACAA GCGCTGACCT GCACAATGAC 1200
TGCACAGAGA CCCACACAGG CACCTCGGCT TCAGCACCCC TGGCTGCTGG 1250

TATCTTTGCT CTGGCCTTGG AGGCAAACCC AAATCTTACC TGGAGAGATA 1300
 TGCAGCATCT GGTGTGTCTGG ACCTCTGAGT ACGACCCATT GGCCAGTAAC 1350
 CCAGGTTGGA AAAAGAATGG GGCAGGCTTG ATGGTGAACA GCCGATTGG 1400
 ATTTGGCTTG CTAAATGCCA AAGCTCTGGT GGATTGGCT GATCCTCGGA 1450
 CCTGGAGAAA TGTGCCTGAG AAGAAAGAAT GTGTTGTAAA AGACAATAAC 1500
 TTTGAGCCTA GAGCCCTGAA AGCTAATGGA GAAGTAATTG TTGAAATCCC 1550
 AACAAGAGCT TGTGAAGGAC AAGAAAATGC TATCAAGTCT CTGGAACATG 1600
 TGCAATTTGA AGCAACAATT GAATATTCTC GTAGAGGAGA CCTTCATGTC 1650
 ACACTCACTT CTGCTGTTGG AACCAGCACT GTACTGTTGG CTGAAAGGGA 1700
 AAGAGATACA TCCCCAATG GCTTTAAGAA TTGGGACTTC ATGTCTGTTC 1750
 ATACATGGGG AGAGAATCCT GTAGGCACCT GGACATTGAA AATTACAGAC 1800
 ATGTCTGGAA GAATGCAAAA TGAAGGAAGG ATTGTGAACT GGAAGTTGAT 1850
 TTTGCATGGG ACATCTTCTC AACCAGAGCA CATGAAGCAG CCCC GTGTGT 1900
 ACACATCCTA CAATACAGTC CAGAATGACA GGAGAGGAGT GGAAAAGATG 1950
 GTGAATGTTG TGGAGAAGCG GCCCACACAA AAGAGCCTGA ATGGCAATCT 2000
 CCTGGTACCC AAAA ACTCCA GCAGCAGCAA TGTGGAGGGT AGAAGGGATG 2050
 AGCAGGTACA AGGA ACTCCT TCAAAGGCCA TGCTGCGACT CCTACAAAGT 2100
 GCTTTTAGCA AGAATGCACT TTCAAACAA TCACCAAAGA AGTCTCCAAG 2150
 TGCAAAGCTC AGCATCCCTT ATGAAAGTTT CTATGAAGCC TTGGAAAAGC 2200
 TTAACAAGCC CTCCAAGCTT GAAGGCTCTG AAGACAGTCT GTACAGTGAC 2250
 TATGTTGATG TATTCTATAA CACAAAACCT TATAAGCATA GAGATGACAG 2300
 GCTGCTGCAA GCTCTCATGG ACATCCTAAA TGAGGAGAAT TAAAATAAGG 2350
 AGCTC 2355

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2012 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCTAGATGCA TCTTCCCTCT TCGTCCCCTG CTCCACCACC CTGCGCGCCT 50
CACAGCCCCG CTTTTCACTC CCAAAGAAGG ATGGAGGGCG GTTGTGGATC 100
CCAGTGAAG GCGGCCGGGT TCCTCTTCTG TGTGATGGTT TTTGCGTCTG 150
CCGAGAGACC CGTCTTCACG AATCATTTTC TTGTGGAGTT GCATAAAGAC 200
GGAGAGGAAG AGGCTCGCCA AGTTGCAGCA GAACACGGCT TTGGAGTCCG 250
AAAGCTCCCC TTTGCAGAAG GCCTGTATCA CTTTTATCAC AATGGGCTTG 300
CAAAGGCCAA AAGAAGACGC AGCCTACACC ATAAGCGGCA GCTAGAGAGA 350
GACCCAGGA TAAAGATGGC GCTGCAACAA GAAGGATTTG ACCGTAAAAA 400
GAGAGGGTAC AGGGACATCA ATGAGATTGA CATCAACATG AATGATCCTC 450
TCTTTACAAA GCAATGGTAC CTGTTCAACA CTGGGCAAGC CGATGGAAC 500
CCTGGGCTAG ACTTGAACGT GGCCGAAGCC TGGGAGCTGG GATACACAGG 550
AAAAGGAGTG ACCATTGGAA TTATGGATGA TGAATTGAC TATCTCCACC 600
CAGACCTGGC CTACAACTAC AACGCTGATG CAAGTTATGA CTTCAGCAGC 650
AATGACCCCT ACCCATACCC TCGATACACA GATGACTGGT TCAACAGCCA 700
TGGAAGTAGG TGTGCAGGAG AAGTTTCTGC TGCAGCCAGC AACAAATATCT 750
GTGGAGTCGG CGTAGCATAA AACTCCAAGG TGGCAGGGAT CCGGATGCTG 800
GACCAGCCCT TTATGACAGA CATCATCGAA GCCTCCTCCA TCAGCCACAT 850
GCCTCAACTG ATCGACATCT ACAGTGCAAG CTGGGGCCCC ACAGACAATG 900
GGAAGACGGT TGATGGGCCC CGAGAGCTCA CACTCCAGGC CATGGCTGAT 950
GGCGTGAACA AGGGCCGTGG GGGCAAAGGC AGCATCTATG TGTGGGCCTC 1000
TGGGGACGGT GGCAGCTACG ATGACTGCAA CTGTGACGGC TATGCTTCAA 1050
GCATGTGGAC CATCTCCATC AACTCAGCCA TCAATGATGG CAGGACTGCC 1100
TTGTATGATG AGAGTTGCTC TTCCACCTTA GCATCCACCT TCAGCAATGG 1150
GAGGAAGAGG AATCCTGAGG CTGGTGTGGC TACCACAGAC TTGTATGGCA 1200
ACTGTACTCT GAGACACTCT GGGACATCTG CAGCTGCTCC GGAGGCAGCT 1250
GGCGTGTTTG CATTAGCTTT GGAGGCTAAC CTGGATCTGA CCTGGAGAGA 1300
CATGCAACAT CTGACTGTGC TCACCTCCAA GCGGAACCAG CTTCATGATG 1350
AGGTTTCATCA GTGGCGACGG AATGGGGTTG GCCTGGAATT TAATCACCTC 1400
TTTGGCTACG GAGTCCTTGA TGCAGGTGCC ATGGTGAAAA TGGCTAAAGA 1450

CTGGAAACT GTTCCGAGA GATTCCATTG TGTGGGAGGC TCTGTGCAGA 1500
 ACCCTGAAAA AATACCACCC ACCGGCAAGC TGGTACTGAC CCTCAAAACA 1550
 AATGCATGTG AGGGGAAGGA AACTTCGTC CGCTACCTGG AGCACGTCCA 1600
 AGCTGTCATC ACAGTCAACG CGACCAGGAG AGGAGACCTG AACATCAACA 1650
 TGACCTCCCC AATGGGCACC AAGTCCATTT TGCTGAGCCG GCGTCCCAGA 1700
 GACGACGACT CCAAGGTGGG CTTTGACAAG TGGCCTTTCA TGACCACCCA 1750
 CACCTGGGGG GAGGATGCCC GAGGGACCTG GACCCTGGAG CTGGGGTTTG 1800
 TGGGCAGTGC ACCACAGAAG GGGTTGCTGA AGGAATGGAC CCTGATGCTT 1850
 CACGGCACAC AGAGCGCCCC ATACATCGAT CAGGTGGTGA GGGATTACCA 1900
 GTCTAAGCTG GCCATGTCCA AGAAGCAGGA GCTGGAGGAA GAGCTGGATG 1950
 AGGCTGTGGA GAGAAGCCTG CAAAGTATCC TGAGAAAGAA CTAGGGCCAC 2000
 GCTTCCGAAT TC 2012

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 753 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Glu	Gln	Arg	Gly	Trp	Thr	Leu	Gln	Cys	Thr	Ala	Phe	Ala	Phe	1	5	10	15
Phe	Cys	Val	Trp	Cys	Ala	Leu	Ser	Ser	Val	Lys	Ala	Lys	Arg	Gln	20	25	30	
Phe	Val	Asn	Glu	Trp	Ala	Ala	Glu	Ile	Pro	Gly	Gly	Gln	Glu	Ala	35	40	45	
Ala	Ser	Ala	Ile	Ala	Glu	Glu	Leu	Gly	Tyr	Asp	Leu	Leu	Gly	Gln	50	55	60	
Ile	Gly	Ser	Leu	Glu	Asn	His	Tyr	Leu	Phe	Lys	His	Lys	Ser	His	65	70	75	
Pro	Arg	Arg	Ser	Arg	Arg	Ser	Ala	Leu	His	Ile	Thr	Lys	Arg	Leu	80	85	90	
Ser	Asp	Asp	Asp	Arg	Val	Thr	Trp	Ala	Glu	Gln	Gln	Tyr	Glu	Lys	95	100	105	
Glu	Arg	Ser	Lys	Arg	Ser	Val	Gln	Lys	Asp	Ser	Ala	Leu	Asp	Leu	110	115	120	

Phe Asn Asp Pro Met Trp Asn Gln Gln Trp Tyr Leu Gln Asp Thr	125	130	135
Arg Met Thr Ala Ala Leu Pro Lys Leu Asp Leu His Val Ile Pro	140	145	150
Val Trp Glu Lys Gly Ile Thr Gly Lys Gly Val Val Ile Thr Val	155	160	165
Leu Asp Asp Gly Leu Glu Trp Asn His Thr Asp Ile Tyr Ala Asn	170	175	180
Tyr Asp Pro Glu Ala Ser Tyr Asp Phe Asn Asp Asn Asp His Asp	185	190	195
Pro Phe Pro Arg Tyr Asp Leu Thr Asn Glu Asn Lys His Gly Thr	200	205	210
Arg Cys Ala Gly Glu Ile Ala Met Gln Ala Asn Asn His Lys Cys	215	220	225
Gly Val Gly Val Ala Tyr Asn Ser Lys Val Gly Gly Ile Arg Met	230	235	240
Leu Asp Gly Ile Val Thr Asp Ala Ile Glu Ala Ser Ser Ile Gly	245	250	255
Phe Asn Pro Gly His Val Asp Ile Tyr Ser Ala Ser Trp Gly Pro	260	265	270
Asn Asp Asp Gly Lys Thr Val Glu Gly Pro Gly Arg Leu Ala Gln	275	280	285
Lys Ala Phe Glu Tyr Gly Val Lys Gln Gly Arg Gln Gly Lys Gly	290	295	300
Ser Ile Phe Val Trp Ala Ser Gly Asn Gly Gly Arg Gln Gly Asp	305	310	315
Asn Cys Asp Cys Asp Gly Tyr Thr Asp Ser Ile Tyr Thr Ile Ser	320	325	330
Ile Ser Ser Ala Ser Gln Gln Gly Leu Ser Pro Trp Tyr Ala Glu	335	340	345
Lys Cys Ser Ser Thr Leu Ala Thr Ser Tyr Ser Ser Gly Asp Tyr	350	355	360
Thr Asp Gln Arg Ile Thr Ser Ala Asp Leu His Asn Asp Cys Thr	365	370	375
Glu Thr His Thr Gly Thr Ser Ala Ser Ala Pro Leu Ala Ala Gly	380	385	390
Ile Phe Ala Leu Ala Leu Glu Ala Asn Pro Asn Leu Thr Trp Arg	395	400	405

Asp Met Gln His	Leu Val Val Trp Thr	Ser Glu Tyr Asp Pro	Leu
410		415	420
Ala Ser Asn Pro	Gly Trp Lys Lys Asn	Gly Ala Gly Leu Met	Val
425		430	435
Asn Ser Arg Phe	Gly Phe Gly Leu Leu	Asn Ala Lys Ala Leu	Val
440		445	450
Asp Leu Ala Asp	Pro Arg Thr Trp Arg	Asn Val Pro Glu Lys	Lys
455		460	465
Glu Cys Val Val	Lys Asp Asn Asn Phe	Glu Pro Arg Ala Leu	Lys
470		475	480
Ala Asn Gly Glu	Val Ile Val Glu Ile	Pro Thr Arg Ala Cys	Glu
485		490	495
Gly Gln Glu Asn	Ala Ile Lys Ser Leu	Glu His Val Gln Phe	Glu
500		505	510
Ala Thr Ile Glu	Tyr Ser Arg Arg Gly	Asp Leu His Val Thr	Leu
515		520	525
Thr Ser Ala Val	Gly Thr Ser Thr Val	Leu Leu Ala Glu Arg	Glu
530		535	540
Arg Asp Thr Ser	Pro Asn Gly Phe Lys	Asn Trp Asp Phe Met	Ser
545		550	555
Val His Thr Trp	Gly Glu Asn Pro Val	Gly Thr Trp Thr Leu	Lys
560		565	570
Ile Thr Asp Met	Ser Gly Arg Met Gln	Asn Glu Gly Arg Ile	Val
575		580	585
Asn Trp Lys Leu	Ile Leu His Gly Thr	Ser Ser Gln Pro Glu	His
590		595	600
Met Lys Gln Pro	Arg Val Tyr Thr Ser	Tyr Asn Thr Val Gln	Asn
605		610	615
Asp Arg Arg Gly	Val Glu Lys Met Val	Asn Val Val Glu Lys	Arg
620		625	630
Pro Thr Gln Lys	Ser Leu Asn Gly Asn	Leu Leu Val Pro Lys	Asn
635		640	645
Ser Ser Ser Ser	Asn Val Glu Gly Arg	Arg Asp Glu Gln Val	Gln
650		655	660
Gly Thr Pro Ser	Lys Ala Met Leu Arg	Leu Leu Gln Ser Ala	Phe
665		670	675
Ser Lys Asn Ala	Leu Ser Lys Gln Ser	Pro Lys Lys Ser Pro	Ser
680		685	690

Ala Lys Leu Ser Ile Pro Tyr Glu Ser Phe Tyr Glu Ala Leu Glu
695 700 705

Lys Leu Asn Lys Pro Ser Lys Leu Glu Gly Ser Glu Asp Ser Leu
710 715 720

Tyr Ser Asp Tyr Val Asp Val Phe Tyr Asn Thr Lys Pro Tyr Lys
725 730 735

His Arg Asp Asp Arg Leu Leu Gln Ala Leu Met Asp Ile Leu Asn
740 745 750

Glu Glu Asn

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 637 amino acids

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Gly Gly Cys Gly Ser Gln Trp Lys Ala Ala Gly Phe Leu
1 5 10 15

Phe Cys Val Met Val Phe Ala Ser Ala Glu Arg Pro Val Phe Thr
20 25 30

Asn His Phe Leu Val Glu Leu His Lys Asp Gly Glu Glu Glu Ala
35 40 45

Arg Gln Val Ala Ala Glu His Gly Phe Gly Val Arg Lys Leu Pro
50 55 60

Phe Ala Glu Gly Leu Tyr His Phe Tyr His Asn Gly Leu Ala Lys
65 70 75

Ala Lys Arg Arg Arg Ser Leu His His Lys Arg Gln Leu Glu Arg
80 85 90

Asp Pro Arg Ile Lys Met Ala Leu Gln Gln Glu Gly Phe Asp Arg
95 100 105

Lys Lys Arg Gly Tyr Arg Asp Ile Asn Glu Ile Asp Ile Asn Met
110 115 120

Asn Asp Pro Leu Phe Thr Lys Gln Trp Tyr Leu Phe Asn Thr Gly
125 130 135

Gln Ala Asp Gly Thr Pro Gly Leu Asp Leu Asn Val Ala Glu Ala
140 145 150

Trp Glu Leu Gly Tyr Thr Gly Lys Gly Val Thr Ile Gly Ile Met
155 160 165

Asp Asp Gly Ile	Asp Tyr Leu His Pro	Asp Leu Ala Tyr Asn Tyr	170	175	180
Asn Ala Asp Ala	Ser Tyr Asp Phe Ser	Ser Asn Asp Pro Tyr Pro	185	190	195
Tyr Pro Arg Tyr	Thr Asp Asp Trp Phe	Asn Ser His Gly Thr Arg	200	205	210
Cys Ala Gly Glu	Val Ser Ala Ala Ala	Ser Asn Asn Ile Cys Gly	215	220	225
Val Gly Val Ala	Tyr Asn Ser Lys Val	Ala Gly Ile Arg Met Leu	230	235	240
Asp Gln Pro Phe	Met Inr Asp Ile Ile	Glu Ala Ser Ser Ile Ser	245	250	255
His Met Pro Gln	Leu Ile Asp Ile Tyr	Ser Ala Ser Trp Gly Pro	260	265	270
Thr Asp Asn Gly	Lys Thr Val Asp Gly	Pro Arg Glu Leu Thr Leu	275	280	285
Gln Ala Met Ala	Asp Gly Val Asn Lys	Gly Arg Gly Gly Lys Gly	290	295	300
Ser Ile Tyr Val	Trp Ala Ser Gly Asp	Gly Gly Ser Tyr Asp Asp	305	310	315
Cys Asn Cys Asp	Gly Tyr Ala Ser Ser	Met Trp Thr Ile Ser Ile	320	325	330
Asn Ser Ala Ile	Asn Asp Gly Arg Thr	Ala Leu Tyr Asp Glu Ser	335	340	345
Cys Ser Ser Thr	Leu Ala Ser Thr Phe	Ser Asn Gly Arg Lys Arg	350	355	360
Asn Pro Glu Ala	Gly Val Ala Thr Thr	Asp Leu Tyr Gly Asn Cys	365	370	375
Thr Leu Arg His	Ser Gly Thr Ser Ala	Ala Ala Pro Glu Ala Ala	380	385	390
Gly Val Phe Ala	Leu Ala Leu Glu Ala	Asn Leu Asp Leu Thr Trp	395	400	405
Arg Asp Met Gln	His Leu Thr Val Leu	Thr Ser Lys Arg Asn Gln	410	415	420
Leu His Asp Glu	Val His Gln Trp Arg	Arg Asn Gly Val Gly Leu	425	430	435
Glu Phe Asn His	Leu Phe Gly Tyr Gly	Val Leu Asp Ala Gly Ala	440	445	450

Met Val Lys Met Ala Lys Asp Trp Lys Thr Val Pro Glu Arg Phe	455	460	465
His Cys Val Gly Gly Ser Val Gln Asn Pro Glu Lys Ile Pro Pro	470	475	480
Thr Gly Lys Leu Val Leu Thr Leu Lys Thr Asn Ala Cys Glu Gly	485	490	495
Lys Glu Asn Phe Val Arg Tyr Leu Glu His Val Gln Ala Val Ile	500	505	510
Thr Val Asn Ala Thr Arg Arg Gly Asp Leu Asn Ile Asn Met Thr	515	520	525
Ser Pro Met Gly Thr Lys Ser Ile Leu Leu Ser Arg Arg Pro Arg	530	535	540
Asp Asp Asp Ser Lys Val Gly Phe Asp Lys Trp Pro Phe Met Thr	545	550	555
Thr His Thr Trp Gly Glu Asp Ala Arg Gly Thr Trp Thr Leu Glu	560	565	570
Leu Gly Phe Val Gly Ser Ala Pro Gln Lys Gly Leu Leu Lys Glu	575	580	585
Trp Thr Leu Met Leu His Gly Thr Gln Ser Ala Pro Tyr Ile Asp	590	595	600
Gln Val Val Arg Asp Tyr Gln Ser Lys Leu Ala Met Ser Lys Lys	605	610	615
Gln Glu Leu Glu Glu Glu Leu Asp Glu Ala Val Glu Arg Ser Leu	620	625	630
Gln Ser Ile Leu Arg Lys Asn	635		

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCAAAATCTA GAYKGCNATY GTNGAYGAKG GN 32

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs

(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAGCATGAGC TCNGGRGCRG CRGCNGANCC 30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATATCACTC AGATCGATGA ATTCGAGCTC 30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTTCTA GAGGATCCCT CTGGTGGATT TGG 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTTGAAT TCTCCAACCC CACACTTGTG 30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGATCGATGA ATTCGAGCTC 20

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CATTCTCGAA AAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CATTCTGTAA AAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CATTCTAGAG CAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CATTCTAGAA AAGCAAGACA A 21

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CATTCTAGAA AAAGAGCACA A 21

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCTGGAGCA AAGCTTCTCT G 21

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACCTGGAGCG CTAGGTCTCT G 21

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CATAAGCTTA CCATGGCCCT GTGGATGCGC 30

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CATTCTAGAC TAGTTGCAGT AGTTCCTCCAG 30

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Lys Thr Arg Arg
1

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Leu Gln Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Arg Thr Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Arg Gln Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Lys Thr Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CTCTGCCTCC CGCTTGGTCC TGGGTGTGTA G 31

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CACGCTTCTG CCGGGATCCC TC 22

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTCTGCCTCC CGCTTGGTCT TCGGTGTGTA G 31

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GATATGAAGA GCAGATCTTT TGGACCTCCG AGGATG 36

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTTATGGTGT AAGCTTCGTT TTGCTCTGGC CTTTGCAAG 39

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TACAACTCAC CGCGGGTCCT G 21

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGATGGGAT GGGATGATGA CCGTTTCCGC CTTGATGT 38

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ACATCACGGC GGAAACGGTC ATCATCCCAT CCCATCTT 38

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GATATAAGCT TGAGAGTGTA GAAGGGGC 28

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGCTTGACAT CATTGGCTGA CACTTTCGAA CACATGATAG 40

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 41 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AAGATGGGAT GGGATGATGA GCGCCGGACC CTCATGGACA T 41

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 41 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATGTCCATGA GGGTCCGGCG CTCATCATCC CATCCCATCT T 41

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GATATAAGCT TGAGAGTGTA GAAGGGGC 28

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGAGCGACAG CACCCCCTTG GAGCCCCCGC CCTTGATCT CATGGAGGAT 50

T 51

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCAGTCCCC TCGTCGGGCG GGGTCCGAGT GCCGTTTCCG CCGTGATGTT 50

C 51

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ACGTGGGCAG CCCCGTGGTG GCGAACAGAA CATCACGGCG GAAACGGC 48

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TGTTCCGCCAC CACGGGGCTG CCCACGTAAT CCTCCATGAG ATACAAGG 48

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTCGGACCCC GCCCGACGAG GGGAGC 26

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCGGGGGCTC CAAGGGGGTG CTGTCGC 27

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GTAGAACAAC ATGGACATGG TGGCAATATT GTCGACTCTG GAGTCGACCT 50

GCAG 54

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CACATAAAAC AAGATGGACA TGGTCTTGTT CACCTGTAGG ATCCCCGG 48

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AGTAAGGAAA AGGATGGTCA TGGTGGAGGT CGACAAGCTT GAGAATTCAA 50

TCG 53

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Xaa Xaa Xaa Arg

1

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

His Ser Arg Lys Lys Arg Gln

1

5

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: Amino Acid
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

His Ser Val Lys Lys Arg Gln

1

5

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

His Ser Arg Ala Lys Arg Gln
1 5

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

His Ser Arg Lys Ala Arg Gln
1 5

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

His Ser Arg Lys Arg Ala Gln
1 5

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Thr Trp Ser Lys Ala Ser Gln
1 5

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Thr Trp Ser Ala Arg Ser Gln
1 5

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CATAAGCTTA CCATGGCCCT GTGGATGCGC 30

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CATTCTAGAC TAGTTGCAGT AGTTCTCCAG 30

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GCTTCCACCA GGTCCGATCC GCACAGGTG 29